

Evaluation of a bespoke training to increase uptake by midwifery teams of NICE Guidance for membrane sweeping to reduce induction of labour:

Kenyon, Sara; Dann, Sophie; Hope, Lucy; Clarke, Paula ; Hogan, Amanda; Jenkinson, David; Hemming, Karla

DOI:

[10.1186/s13063-017-2106-1](https://doi.org/10.1186/s13063-017-2106-1)

License:

Creative Commons: Attribution (CC BY)

Document Version

Peer reviewed version

Citation for published version (Harvard):

Kenyon, S, Dann, S, Hope, L, Clarke, P, Hogan, A, Jenkinson, D & Hemming, K 2017, 'Evaluation of a bespoke training to increase uptake by midwifery teams of NICE Guidance for membrane sweeping to reduce induction of labour: a stepped wedge cluster randomised design', *Trials*, vol. 18, no. 357. <https://doi.org/10.1186/s13063-017-2106-1>

[Link to publication on Research at Birmingham portal](#)

General rights

Unless a licence is specified above, all rights (including copyright and moral rights) in this document are retained by the authors and/or the copyright holders. The express permission of the copyright holder must be obtained for any use of this material other than for purposes permitted by law.

- Users may freely distribute the URL that is used to identify this publication.
- Users may download and/or print one copy of the publication from the University of Birmingham research portal for the purpose of private study or non-commercial research.
- User may use extracts from the document in line with the concept of 'fair dealing' under the Copyright, Designs and Patents Act 1988 (?)
- Users may not further distribute the material nor use it for the purposes of commercial gain.

Where a licence is displayed above, please note the terms and conditions of the licence govern your use of this document.

When citing, please reference the published version.

Take down policy

While the University of Birmingham exercises care and attention in making items available there are rare occasions when an item has been uploaded in error or has been deemed to be commercially or otherwise sensitive.

If you believe that this is the case for this document, please contact UBIRA@lists.bham.ac.uk providing details and we will remove access to the work immediately and investigate.

Trials

Evaluation of a bespoke training to increase uptake by midwifery teams of NICE Guidance for membrane sweeping to reduce induction of labour: a stepped wedge cluster randomised design --Manuscript Draft--

Manuscript Number:	TRLS-D-16-00735R3	
Full Title:	Evaluation of a bespoke training to increase uptake by midwifery teams of NICE Guidance for membrane sweeping to reduce induction of labour: a stepped wedge cluster randomised design	
Article Type:	Research	
Funding Information:	National Institute for Health Research (CLAHRC 2013 10018)	Not applicable
Abstract:	<p>Background National guidance recommends pregnant women are offered membrane sweeping at term to reduce induction of labour. Local audit suggested this was not being undertaken routinely across two maternity units in the West Midlands, UK between March and November 2012.</p> <p>Methods Bespoke training session for midwifery teams (9 community and 1 antenatal clinic) was developed to address identified barriers to encourage offer of membrane sweeping, together with an information leaflet for women and appointment of a champion within each team. The timing of training session on membrane sweeping to 10 midwifery teams was randomly allocated using a stepped wedge cluster randomised design. All women who gave birth in the Trusts after 39+3/40 weeks gestation within the study time period were eligible. Relevant anonymised data were extracted from maternity notes for three months before and after training. Data were analysed using a generalised linear mixed model, allowing for clustering and adjusting for temporal effects.</p> <p>Main outcomes Primary outcomes were number of women offered and accepting membrane sweeping and average number of sweeps per woman. Sub-group comparisons were undertaken for adherence to Trust guidance and potential influence of pre-specified maternal characteristics. Data included whether sweeping was offered but declined and no record of membrane sweeping.</p> <p>Results Training was given to all teams as planned. Analyses included data from 2,787 of the 2,864 (97%) eligible low risk women over 39+4 weeks pregnant. Characteristics of the women were similar before and after training. No evidence of difference in proportion of women being offered and accepting membrane sweeping (44.4% before training versus 46.8% after training (adjusted Relative risk (aRR) 0.90, 95% CI (0.71, 1.13)), nor in average number of sweeps per woman (0.603 versus 0.627 (aRR) 0.83. 95% CI (0.67, 1.01)). No differences in any secondary outcomes nor influence of maternal characteristics was demonstrated. The midwives evaluated training positively.</p> <p>Conclusions This stepped wedge cluster trial enabled randomised evaluation within a natural roll out and demonstrates the importance of robust evaluation in circumstances in which it is rarely undertaken. While the midwives evaluated the training positively, it did not appear to change practice.</p> <p>Retrospective Trials registration: Registered 23.08.2016 Biomed Central ISRCTN 14300475 Key words: stepped wedge cluster randomised evaluation of training for community midwives</p>	
Corresponding Author:	Sara Kenyon University of Birmingham College of Medical and Dental Sciences Birmingham, UNITED KINGDOM	
Corresponding Author Secondary Information:		

Corresponding Author's Institution:	University of Birmingham College of Medical and Dental Sciences
Corresponding Author's Secondary Institution:	
First Author:	Sara Kenyon
First Author Secondary Information:	
Order of Authors:	Sara Kenyon
	Sophie Dann
	Lucy Hope
	Paula Clarke
	Amanda Hogan
	David Jenkinson
	Karla Hemming
Order of Authors Secondary Information:	
Response to Reviewers:	<p>Institute of Applied Health Research Edgbaston Campus University of Birmingham Edgbaston B15 2TT</p> <p>BioMed Central Floor 6, 236 Gray's Inn Road London WC1X 8HB United Kingdom Dear Editor, TRLS-D-16-00735 Evaluation of a bespoke training to increase uptake by midwifery teams of NICE Guidance for membrane sweeping to reduce induction of labour: a stepped wedge cluster randomised design</p> <p>We are resubmitting this response as there seems to be some problem with your server. We are most grateful for the opportunity to respond to the additional reviewers comments and have done so in the table below. We hope that you will review these and the revised paper and consider the paper suitable for publication.</p> <p>Reviewer #3: Reviewer's comments This study assessed the effect of bespoke training on uptake of membrane sweeping by the midwifery teams to reduce induction of labour using stepped wedge cluster randomized design. However this manuscript will require major revision before considering it for publication</p> <p>Thank you for taking the time to review this paper. We hope that our response and revisions will mean you do consider it suitable for publication.</p> <p>Abstract: 1)The authors should restructure the abstract section in-line with this format: Background, Methods, Results, and Conclusions</p> <p>The abstract we have submitted is in line with the format you describe and so no change has been made.</p> <p>Methods 2)The authors claimed that they used stepped wedge cluster randomized design to assess the outcome. Typically in stepped wedge cluster randomized design the recruited clusters will randomly crossover from pre-intervention phase to intervention phase. It indicates that the duration of pre-intervention and intervention phase for each</p>

cluster will not be the same except for cluster that crossed over at the middle of trial. Clusters that crossed over to the intervention phase early in the trial should have a longer intervention phase than pre-intervention phase. In addition, at the early part of a stepped wedge cluster randomized study most of the clusters will be in the pre-intervention phase while at the later part of the study most the clusters would have crossed over to the intervention phase.(1,2) Thus, this paper did not apply "stepped wedge cluster randomized design" as they claimed rather the authors applied "before and after design"

Thank you for highlighting that there is no universally accepted definition of the SW-CRT, but the paper which the reviewer cites (Hemming et al BMJ) uses the following broad definition "The design involves random and sequential crossover of clusters from control to intervention until all clusters are exposed." Clearly the study reported here does meet the broad definition of the SW-CRT used by the paper the reviewer cited. This is not a before and after design, as we have data on which to estimate temporal trends -and the analysis is adjusted for this. This is of course the key feature of the SW-CRT which means it provides a much higher level of evidence than a simple pre and post study.

3)The authors considered two primary outcomes (proportion of women offered and accepting membrane sweeping and average number of sweeps per woman) in this study. It is important to note that intervention study should not have more than one primary outcome but the outcome can either be a composite outcome or single outcome. The reason provided by the authors that the two outcomes are highly correlated will not justify why the authors considered them as primary outcomes. Moreover the two outcomes are not the same; one is a binary outcome while the other is a continuous outcome. This issue should be addressed

It is not uncommon for studies to have two primary outcomes and as these were pre-specified in the trial protocol they cannot be changed. We have interpreted our findings cautiously - and what is more this study finds no evidence of any effect.

4) The authors should justify why they considered number of women offered and accepting membrane sweeping and average number of sweeps per woman instead of considering the proportion of ELIGIBLE women who were offered membrane sweeping and accepting membrane sweeping.

We have included all the women eligible for membrane sweeping and report outcomes for those women. We have revisited the abstract to make this point clearer, as we believe it was within the actual paper.

5)The authors mentioned in their response to the previous reviewer that the primary outcome was "proportion of women offered and accepting membrane sweeping and average number of sweeps per woman" however in the abstract they reported "number of women offered and accepting membrane sweeping and average number of sweeps per woman" as primary outcome. The authors should note that the primary outcome of this study should be "the proportion of ELIGIBLE women who were offered and accepting membrane sweeping"

We report exclusions clearly both in the text and the Consort Flow diagram. The study included all women eligible for membrane sweeping and we could not report outcomes for a very small number (2-3%).

Minor point - can be fixed

6)The authors stated that they followed Hussey and Hughes to estimate the sample size but they did not the intervention effect that they used. Please address this important issue.

We did follow the H and H approach. The intervention effect is reported in the paper in the first two paragraphs on page 9.

7)The authors stated that they used mixed effects Poisson regression model to examined the difference in the proportion of women being swept in before and after the training. I will like to know why the authors considered apply Poisson regression when the outcome is not a count variable.

Poisson regression is an accepted method to estimate relative risks when the outcome is binary. We use robust standard errors to account for the mis-specification of the variances. By oversight we did not mention that we used robust standard errors and this has now been added.

8)The authors claimed that they used mixed effects Poisson regression model but did not report the random effect part of the model in the result table. Kindly provide the results.

We reported ICCs on the proportions scale (as opposed to variances on the log scale).
9)The authors performed sub-group analysis but did not consider it when estimating the sample size. Kindly address this issue.

It is common practice to pre-specify a small number of subgroups without being fully powered for interactions, and this is what we have done.

References

1. Hemming K, Haines TP, Chilton PJ, Girling a. J, Lilford RJ. The stepped wedge cluster randomised trial: rationale, design, analysis, and reporting. Bmj [Internet]. 2015;350(feb06 1):h391-h391. Available from: <http://www.bmj.com/cgi/doi/10.1136/bmj.h391>
2. Hussey MA, Hughes JP. Design and analysis of stepped wedge cluster randomized trials. Contemp Clin Trials. 2007;28(2):182-91.

We would be very grateful for your consideration and look forward to hearing from you.
Yours faithfully,

Sara Kenyon
Professor of Evidence Based Maternity Care

[Click here to view linked References](#)

Evaluation of a bespoke training to increase uptake by midwifery teams of NICE Guidance for membrane sweeping to reduce induction of labour: a stepped wedge cluster randomised design

Sara Kenyon, Institute of Applied Health Research, College of Medical and Dental Sciences, University of Birmingham, Edgbaston, Birmingham, B15 2TT, UK (s.kenyon@bham.ac.uk).

Corresponding author

Sophie Dann, Institute of Applied Health Research, College of Medical and Dental Sciences, University of Birmingham, Edgbaston, Birmingham, B15 2TT, UK (sophiedann@gmail.com)

Lucy Hope, Institute of Health & Society, University of Worcester, Henwick Grove, WR2 6AJ (l.hope@worc.ac.uk)

Paula Clarke, Birmingham Women's NHS Foundation Trust, Mindelsohn Way, Birmingham, West Midlands B15 2TG, UK (paula.clarke@bwnft.nhs.uk)

Amanda Hogan, Quality Assurance Team for Adult and Ante-natal/Newborn Screening Programmes, Midlands and East Region, 1st Floor, 5 St Philip's Place, Birmingham B3 2PW, UK (amandahogan1@nhs.net)

David Jenkinson, Division of Health Sciences, Warwick Medical School, University of Warwick, Coventry, CV4 7AL, UK (D.J.Jenkinson@warwick.ac.uk)

Karla Hemming, Institute of Applied Health Research, College of Medical and Dental Sciences, University of Birmingham, Edgbaston, Birmingham, B15 2TT, UK (k.hemming@bham.ac.uk)

ABSTRACT

Background

National guidance recommends pregnant women are offered membrane sweeping at term to reduce induction of labour. Local audit suggested this was not being undertaken routinely across two maternity units in the West Midlands, UK between March and November 2012.

Methods

Bespoke training session for midwifery teams (9 community and 1 antenatal clinic) was developed to address identified barriers to encourage offer of membrane sweeping, together with an information leaflet for women and appointment of a champion within each team.

The timing of training session on membrane sweeping to 10 midwifery teams was randomly allocated using a stepped wedge cluster randomised design. All women who gave birth in the Trusts after 39+3/40 weeks gestation within the study time period were eligible. Relevant anonymised data were extracted from maternity notes for three months before and after training. Data were analysed using a generalised linear mixed model, allowing for clustering and adjusting for temporal effects.

Main outcomes

Primary outcomes were number of women offered and accepting membrane sweeping and average number of sweeps per woman. Sub-group comparisons were undertaken for adherence to Trust guidance and potential influence of pre-specified maternal characteristics. Data included whether sweeping was offered but declined and no record of membrane sweeping.

Results

Training was given to all teams as planned. Analyses included data from 2,787 of the 2,864 (97%) eligible low risk women over 39+4 weeks pregnant. Characteristics of the women were similar before and after training. No evidence of difference in proportion of women being offered and accepting membrane sweeping (44.4% before training versus 46.8% after training (adjusted Relative risk (aRR) 0.90, 95% CI (0.71, 1.13)), nor in average number of sweeps per woman (0.603 versus 0.627 (aRR) 0.83. 95% CI (0.67, 1.01)). No differences in any secondary outcomes nor influence of maternal characteristics was demonstrated. The midwives evaluated training positively.

Conclusions

This stepped wedge cluster trial enabled randomised evaluation within a natural roll out and demonstrates the importance of robust evaluation in circumstances in which it is rarely undertaken. While the midwives evaluated the training positively, it did not appear to change practice.

Retrospective Trials registration: Registered 23.08.2016 Biomed Central ~~ISRCTN14300475~~

[ISRCTN14300475](#)

Key words: stepped wedge cluster randomised evaluation of training for community midwives

Background

In 2010-2011, approximately 21% of births in the UK were induced (NHS Maternity Statistics [1]). Induction is undertaken for a variety of indications, with post-term pregnancy being one of the most common. Induction of labour can have a negative impact on women's birth experiences and is found by women to be more painful than spontaneous labour [2]. At the time the study was planned (2010-2011) women who labour spontaneously had a Caesarean section (CS) rate of 11% and instrumental birth rate of 13%. In contrast, women who were induced had higher rates of CS (22%) and instrumental birth (17%). It is therefore important to do as much as possible to reduce the numbers of women requiring induction.

The National Institute for Health and Care Excellence (NICE) Guidance on Inducing Labour [3] reviewed the evidence on various relatively non-invasive methods of inducing labour, namely membrane sweeping, herbal supplements, acupuncture, homeopathy, castor oil, hot baths and enemas, sexual intercourse and breast stimulation. They concluded that there was insufficient evidence to recommend any of them other than membrane sweeping, which they recommend is undertaken to reduce induction of labour. See Box 1 for recommendations regarding membrane sweeping from NICE Inducing Labour Guideline. This Guideline was originally published in 2008 and regular Evidence Updates have not found need to update recommendations on the basis of published research, so the Guideline remains current and is due for review again in 2016 [4]. Membrane sweeping is also a NICE Antenatal Care Quality Standard [5].

Box 1. Recommendations regarding membrane sweeping from NICE Inducing Labour Guideline.

Membrane sweeping involves the examining finger passing through the cervix to rotate against the wall of the uterus, to separate the chorionic membrane from the decidua. If the cervix will not admit a finger, massaging around the cervix in the vaginal fornices may achieve a similar effect. For the purpose of this guideline, membrane sweeping is regarded as an adjunct to induction of labour rather than an actual method of induction.

The Bishop score is a group of measurements made by doing a vaginal examination, and is based on the station, dilation, effacement (or length), position and consistency of the cervix. A score of eight

or more generally indicates that the cervix is ripe, or 'favourable' – when there is a high chance of spontaneous labour, or response to interventions made to induce labour.

1.3.1 Membrane sweeping

1.3.1.1 Prior to formal induction of labour, women should be offered a vaginal examination for membrane sweeping.

1.3.1.2 At the 40 and 41 week antenatal visits, nulliparous women should be offered a vaginal examination for membrane sweeping.

1.3.1.3 At the 41 week antenatal visit, parous women should be offered a vaginal examination for membrane sweeping.

1.3.1.4 When a vaginal examination is carried out to assess the cervix, the opportunity should be taken to offer the woman a membrane sweep.

1.3.1.5 Additional membrane sweeping may be offered if labour does not start spontaneously

Assessment of current practice

An audit at Birmingham Women's NHS Foundation Trust (BWNFT) had suggested that not all eligible women were offered membrane sweeping, and maternity managers at Birmingham Heartlands Hospital (BHH) – part of Heart of England NHS Foundation Trust – also felt membrane sweeping was not done according to NICE Guidance in the maternity unit. Guidelines at BHH reflected those of NICE while BWNFT guidelines recommended sweeping of both nulliparous and multiparous women at 40 and 41 weeks. Collaboration between managers at the two trusts and researchers at the University of Birmingham facilitated a robust evaluation of the effect of bespoke training for midwives to increase membrane sweeping to reduce induction of labour using a stepped wedge cluster randomised trial design.

Methods

A stepped wedge cluster randomised trial was used as it was the intention of the healthcare providers that all midwifery teams (9 community and 1 antenatal clinic) receive the training. It was not possible to implement the training module over all teams concurrently and individual randomisation of midwives could not be used as the intervention was delivered to teams so contamination would be unavoidable. The stepped wedge design that was used is illustrated in Supplementary Figure 1. Data were collected for each team from 12 weeks prior to their team's training being delivered, and concluded at 12 weeks following the team's training. The order in

1 which each team received the training was randomised. In this pragmatic evaluation, we were
2 limited by the number of midwifery teams in the area and so were not able to increase the number
3 of clusters.
4

5
6 Randomisation was performed in Stata, at a single point in time, by the study statistician (KH). Each
7 of the 10 midwifery teams were allocated a unique ID. These 10 unique IDs were then randomly
8 sorted to provide the order in which the teams would be trained. The teams were informed of their
9 allocation date in sequential order once the previous team had set the date for training (a two week
10 period when training should be undertaken). Training took place between May and September 2012,
11 and data were collected from 5th March 2012 to 26th November 2012, as shown in Figure 1.
12
13
14
15
16

17 ***Intervention***

18
19
20 Midwives in the teams (9 community and 1 antenatal clinic) received the following:
21

22
23 A generic interactive training package taking approximately an hour containing evidence, practical
24 tips to support practice, and a leaflet for women. The training was supported by written materials
25 for the midwives.
26
27

28
29 A lead midwife (Champion) was identified in each team to be an expert for clinical queries, and to
30 train and remind staff.
31

32
33 A leaflet was developed for women entitled 'Ways to reduce your need for induction of labour:
34 membrane sweeping', with input from the local Maternity Services Liaison Committee (service user
35 group), with the intention that it would be given to women at 36/40 weeks gestation at their routine
36 antenatal visit.
37
38
39
40

41
42 Training was led by the Practice Development Midwives/Consultant Midwives/Training Department
43 in each Trust, supported by the NIHR Collaboration for Leadership and Applied Health Research and
44 Care in Birmingham and the Black Country (CLAHRC BBC) researchers from the University of
45 Birmingham.
46
47

48
49 As part of the development of the intervention one midwife from each team attended a group
50 facilitated by CLAHRC researchers to discuss interventions they currently suggest to women or use to
51 reduce induction of labour, and to understand any barriers to the midwives sweeping women's
52 membranes. The concerns identified by the midwives included lack of time, unsuitable venue, and
53 the importance of maternal preparation. Some midwives also stated they were unsure of the
54 technique and its effectiveness, and expressed some reluctance to undertake membrane sweeping
55 (painful for women and don't like to interfere with nature).
56
57
58
59
60
61

Inclusion/exclusion criteria

All the midwifery teams were included in Trusts/Units where membrane sweeping was felt not to be done according to NICE and Trust Guidance. The number of clusters (i.e. teams) included was pre-determined based on teams the Trusts had decided to roll-out the training package to.

Eligible women included all those who gave birth over 39+3 weeks at BWNFT or BHH within the study period (March to November 2012). Women were included who gave birth after 39+3 weeks (rather than over 40 weeks) as it is plausible that women were not swept on exactly the correct day (i.e. 40 or 41 weeks), so a sweep from 39+4 to 40+3 was considered a sweep at 40 weeks, and 40+4 to 41+3 was considered a sweep at 41 weeks. Women were excluded if they were from outside the area or had an elective CS.

Outcomes

The primary outcomes were the proportion of women offered and accepting a membrane sweep and the average number of sweeps per woman.

Secondary outcomes included onset of labour, mode of birth, and adherence to Trust guidance. NICE guidance recommends that nulliparous women are offered membrane sweeping at 40 and 41 weeks and multiparous women at 41 weeks and BHH had adopted this. At BWNFT, Trust Guidance recommended all women were swept at 40 and 41 weeks regardless of parity. Information was also collected on sweeps offered but declined, where no record of membrane sweeping was found in the maternity notes, and the location of the sweep (community or hospital). If sweeping was abandoned, the reason was collected (e.g. cervical os closed, unable to reach, unable to sweep).

Planned subgroup analysis included whether the numbers of women having a membrane sweep was influenced by maternal age (<20 and >35), parity (nulliparous and multiparous), ethnicity (from antenatal notes), BMI (Body Mass Index) (<18 and >35), or Index of Multiple Deprivation (IMD) based on post code.

Training was evaluated by the midwives using a questionnaire both immediately and six months afterwards.

Data collection

Women were identified for inclusion in the analysis by the individual Trusts' electronic systems, and were included if they gave birth after 39+3 weeks gestation within the study period. It was planned that the data on sweeping status would form part of the mandatory electronic data collected within

each Trust and that this data would be collected from the antenatal notes by the midwife entering birth outcome data onto the Trusts' electronic systems. Information would be recorded as to whether the woman was offered sweeping, whether she accepted, and at what gestation. However, this did not prove possible within the study time at BWNFT, so pseudo-anonymised data was extracted from the handheld antenatal records of 2,864 women and entered onto a bespoke database. At BHH the mandatory electronic data collection began part way through the study (15th June 2012) and pseudo-anonymised data was transferred electronically to the University of Birmingham after that date. Prior to this, all data was extracted from the handheld maternity notes.

Planned data cross-checking with source notes of the electronically-transferred data on a random 20% of women found inaccuracies in 30% of data, so data on membrane sweeping, ethnicity, BMI, post code, and midwifery team was manually extracted from the handheld maternity notes for all women.

Systems were agreed with the Research and Development (R&D) Departments, which ensured only pseudo-anonymised data were transferred to and stored by the University of Birmingham. Data were extracted by members of the University team holding Research Passports and with permission to do so. All data were given a unique study specific number and only that required was extracted for the agreed analysis. Blinding to intervention period was not possible but personnel extracting data were not aware of individual team training dates.

Sample size

From Hospital Episode Statistics we estimated that there would be approximately 12 births after 40 weeks gestation per week in each team. Birth data was collected for each team from 12 weeks prior to training until 12 weeks following training. Given this fixed sample size, we determined what difference in the primary outcome (proportion of women swept) would be detectable with 80% power. We did not make allowances for the co-primary outcomes as these two outcomes are highly correlated.

The calculation depended upon both the current proportion of women being swept and the magnitude of intra-cluster correlation coefficient (ICC) between the proportions of women swept in each of the midwifery teams. Estimates of ICC would ideally come from other similar studies but, in the absence of such evidence, we were guided by a review of estimates of ICCs which found that their values are typically between 0.02 and 0.1 [6]. A small audit suggested that of those eligible for sweeping, 32% of nulliparous women and 57% of multiparous women were currently being swept.

Methods described in Hussey and Hughes [7], [8] were followed to determine power, and implemented using the Stata function [9]. It was estimated that at 5% significance (two tailed) and 80% power, for ICCs between 0.02 and 0.1, and for baseline event rates of between 20% and 60%, the study would have power to detect around a 10% absolute increase in proportion of women being swept. This was an increase felt to be clinically worthwhile.

Analysis

The participants' characteristics were summarised using appropriate summary statistics, grouping them by whether they gave birth before or after the training session. These characteristics included the woman's parity, ethnicity, BMI, and IMD based on post code as well as the Trust caring for her. Teams were classified as being exposed to the intervention the week after the team underwent the training, and births during these transition weeks were not included. The trial was well balanced on all characteristics (Table 1) and so no adjustment was made for patient level characteristics in the outcome analysis.

The primary aim of the study was to evaluate whether there was a difference in the proportion of women being swept in the 12 week period before and after the training session (intervention). To this end, we fitted a mixed effects Poisson regression model, using robust standard errors to account for the misspecification of the variances¹⁰. ~~We included~~including, as explanatory variables, the treatment exposure (before or after training, as a fixed effect), the midwifery team (as a random effect, accounting for the clustering), and calendar time (as a fixed effect). The treatment effect is reported as the adjusted relative risk (aRR) of being offered and accepting a sweep. The other primary outcome (number of sweeps) was also analysed by a mixed effects Poisson regression model, with the same explanatory variables. The treatment effect is the adjusted incidence rate ratio (aIRR) of having one extra sweep after the intervention compared to before it. The secondary outcomes were binary and were also analysed using Poisson models and, again, reporting relative risks. For the analysis of subgroups, the same Poisson regression model was applied to a subset of the data containing the participants that belonged to that subgroup.

All analyses were carried out in duplicate, independently, to verify the results (KH and DJ). Results reported were carried out in R, although the independent verification was carried out in Stata 12. Comparisons will be considered significant at the 5% level and so 95% confidence intervals (CI) are reported throughout.

Results

1 Training was given to all the midwifery teams as planned, with the majority of team members being
2 present (73/108 (67%)). There were 10 midwifery teams (9 community and 1 antenatal clinic) that
3 included an average of 10 midwives. The average size of team varied between the Trusts (BWNFT 14
4 and BHH 7).
5
6

7
8 Of the 2,864 women identified by the Trusts as potentially eligible for inclusion, 2,787 were included
9 in the analysis (1,420 women before training and 1,367 after). 34 women before and 43 women
10 after the training were ineligible (Figure 1) as seen in the CONSORT flow diagram. Data was not
11 available for 14 women (3 before the intervention and eleven after) and so they could not be
12 included in the analysis. Membrane sweeping was offered and refused by 6% of women (Table 2).
13
14 The characteristics of the women before and after training were similar (Table 1).
15
16
17

18
19 There was no evidence of any differences in the primary outcome of numbers of women being
20 offered and accepting membrane sweeping before and after training (44.4% versus 46.8% (aRR)
21 0.90, 95% CI (0.71, 1.13)), nor in the average number of membrane sweeps being undertaken per
22 woman (0.603 versus 0.627 (aRR) 0.83, 95% CI (0.67, 1.01)) (Table 2). Relative risks are adjusted for
23 clustering and underlying temporal trends. There was no evidence of any differences in either
24 secondary outcome, onset of labour or mode of birth (Table 2).
25
26
27
28
29
30

31 Trust-specific results for BWNFT found no evidence of differences in the primary outcome of
32 numbers of women being offered and accepting membrane sweeping before and after training
33 (47.4% versus 51.8% (aRR) 0.87, 95% CI (0.67, 1.14)). However, the average number of membrane
34 sweeps being undertaken per woman had significantly decreased (0.660 versus 0.701 (aRR) 0.71,
35 95% CI (0.55, 0.90)) (Table 2). Improvement in adherence to Trust Guidance was not seen between
36 the two periods (Table 2).
37
38
39
40
41

42 Trust-specific results for BHH found no evidence of differences in the primary outcome of numbers
43 of women being offered and accepting membrane sweeping before and after training (38.7% versus
44 37.7% (aRR) 1.17, 95% CI (0.76, 1.81)) nor in the average number of membrane sweeps being
45 undertaken per woman (0.497 versus 0.493 (aRR) 1.32, 95% CI (0.89, 1.95)) (Table 2). Improvement
46 in adherence to Trust guidance was not seen between the two periods (Table 2).
47
48
49
50
51

52 No differences were seen in any other outcome for either Trust.
53
54

55 The comparison of the effect of selected characteristics on the intervention training demonstrated
56 no individual effect of maternal age, parity, ethnicity, BMI or index of multiple deprivation from
57 postcode (Supplementary Table 1).
58
59
60
61

Response rates to the training questionnaires were good: 69/73 (95%) immediately following training and 60/73 (82%) six months after training. Overall evaluation of training showed knowledge of the evidence and current NICE Guidance regarding membrane sweeping was high before training (average 4/5), improved (to 5/5) immediately after training, and reduced slightly (to 4/5) at six months after training. 60% (36/60) of midwives stated that training had changed their practice (Table 3).

A sensitivity analysis investigated the impact of the additional level of clustering (teams nested within trusts) by including trust as a fixed effect but results were not sensitive to this additional clustering (results not included). The discrepancy between the ratio of the two proportions and the relative risk presented in the table (RR=0.9) arises because the relative risk presented in the table is adjusted for time effects. In fact it is also important to note that the ratio of the percentages swept without adjusting for time is not 0.9, but is in fact 1.05 (=intervention percentage/control percentage=46.8/44.4). That is to say, the raw results suggest that on implementation the point estimate of the percentage of women being swept increased (from 44.4% to 46.8%), hence an increased "risk" of being swept. However, in actual fact, in those clusters and time periods yet to be exposed to the intervention there was an underlying secular trend (Figure S3). However, after adjusting for the underlying secular trend, we demonstrate a reversal of the treatment effect. Although of note, all these changes are small and not statistically significant. We did not examine time by treatment interactions as the study was underpowered for this comparison and this analysis had not been pre-specified.

Discussion

The delivery of the bespoke training package to midwifery teams had little effect on the number of women being offered and accepting membrane sweeping. Had this robust evaluation not been undertaken it may well have been felt that the training had been effective and practice had changed due to the positive feedback from the midwives, and evaluations such as this should be encouraged.

Studies such as these highlight the complexity of changing practice. While we did attempt to address the issues highlighted by the midwives as problematic, it is clear that this was not enough to increase the numbers of women being offered and accepting membrane sweeping. As described earlier, some of the concerns identified by the midwives were beyond the scope of the training (such as the venue) but were discussed. While we did identify a sweeping champion within each team, to provide leadership so important in change, there was no mechanism in place to enable data on sweeping to be regularly fed back to teams, which may have been helpful. While a leaflet was developed for

women, there is evidence that where evidence based information is prioritised over women's or healthcare professionals' experiential knowledge, there is potential conflict [11], and it is plausible that not all women received the leaflet as intended.

The training session was delivered at the team meetings and in a way most suited to adult learning with discussion encouraged. The training session was delivered to the majority of team members (67%), although this did vary between the Trusts with 65% attending at BWNFT and 72% at BHH. Midwives who could not attend the team training session were trained by the champions in their team. The average size of the teams was markedly different (BWNFT 14 and BHH seven) and it is interesting to note that the Trust with the larger teams and slightly lower attendance rates demonstrated a higher level of sweeping, although no improvement was seen overall (we did not test whether this was statistically significant). While this attempt to change practice had high quality evidence to underpin it, reducing the rate of induction of labour would not directly affect the community midwives undertaking the sweeping. The direct effect would have been felt by the women themselves and the Labour Wards in the Trusts and it is plausible that this influenced whether the midwives did change their practice.

Strengths and Limitations

This study provides robust evidence of the effect of a bespoke training package on improving implementation of NICE Guidance to undertake membrane sweeping to reduce induction of labour and such evaluations are relatively uncommon. A recent review [12] suggests that there have been advancements in factors that influence training effectiveness and transfer of training but that robust evaluation should be encouraged and evaluation of methods of getting evidence into practice is essential to informing good quality care. One of the main obstacles to evaluating such interventions, or change, is the lack of opportunity to randomise: evidence of effectiveness is not required before implementation and so changes are often instigated before evaluation. Evaluation after instigation can be possible, either by comparison with other providers who have not instigated change, or with service provision before the change occurred, or with both before and control comparisons, but it is well known that this forms lower quality evidence. Stepped wedge randomised trials have been suggested as a pragmatic and appropriate option for evaluation of service delivery type interventions [13]. Whilst this design has been recommended in service evaluations, it is important to ensure appropriate input from an experienced statistician due to the complex nature of the design and ensuing data analysis. Both BWNFT and BHH planned to sequentially roll out a training module and collaboration with the CLAHRC researchers at the University of Birmingham made this evaluation possible. Data collection was relatively complete for the trial, thus increasing reliability

1 and validity. We did not allow for any multiplicity of outcomes in our power calculation. Whilst the
2 primary outcome, sweeping, has been reported in two different ways (number of sweeps and
3 proportion of women swept) these two outcomes are very highly correlated and any multiplicity
4 correction would be highly conservative.
5
6

7 We observed a significant underlying temporal trend in the proportion of women offering a
8 membrane sweep over the duration of the study period. There are a number of possible
9 explanations for this. Contamination between teams is a possibility, but unlikely as the teams did not
10 mix regularly. It is more likely that the very movement that lead local Trust decision makers to
11 initiate this training package, also penetrated down to the teams and the midwives, akin to a rising
12 tide¹⁴.
13
14
15
16
17

18 One limitation of the study is that data regarding membrane sweeping was collected from the
19 hospital notes, as described earlier, and it could be argued that this does not reflect actual practice.
20 While it is possible, it was felt to be very unlikely that a membrane sweep would be undertaken and
21 not recorded in the notes. The data was extracted by the same data clerk, blind to the date of
22 training of the team, and training was given by clinical midwives as to where this would be recorded.
23
24
25
26
27
28

29 A method for characterising and designing behaviour change interventions, which includes the
30 'COM-B' system, has been developed by Michie et al [15], and it may be that use of such a
31 systematic approach would have improved the number of women being offered and accepting
32 membrane sweeping. The 'COM-B' system provides a framework for understanding behaviour with
33 three essential conditions interacting to generate change, which are capability, motivation and
34 opportunity. Capability is defined as the individual's psychological and physical capacity to engage in
35 the activity concerned and it includes having the necessary knowledge and skills. Motivation is
36 defined as all those brain processes that energize and direct behaviour, not just goals and conscious
37 decision-making. It includes habitual processes, emotional responses, as well as analytical decision-
38 making. Opportunity is defined as all the factors that lie outside the individual that make the
39 behaviour possible or prompt it.
40
41
42
43
44
45
46
47
48

49 Achieving and maintaining behaviour change remains challenging and Michie [16] suggests that
50 meeting it requires a systematic method for analysing the target behaviours as a starting point for
51 designing an intervention; selecting interventions most likely to be effective; publishing details of
52 interventions in trial protocols to enable accurate replication and evidence synthesis and drawing on
53 relevant theory to guide both the intervention design and evaluation. Such knowledge and skills may
54
55
56
57
58
59
60
61
62

not be accessible to the majority of healthcare providers. None the less, use of such methods should be encouraged.

Conclusion

Novel ways of evaluating service change to improve uptake of NICE Guidance should be encouraged and use of the stepped wedge design offers a pragmatic and useful methodology in such situations, even if results showed no significant difference in this instance. In the future use of a systematic approach to the development of behaviour change interventions should be encouraged to increase the likelihood of success and results should be fed back to Trusts to further encourage collaboration and change.

Abbreviations

BHH-Birmingham Heartlands Hospital

BWNFT- Birmingham Women's NHS Foundation Trust

CLAHRC BBC- Collaboration for Leadership and Applied Health Research and Care Birmingham and Black Country

ICC- Inter cluster correlation

NHS- National Health Service

NICE- National Institute for Health and Care Excellence

R&D- Research and Development

Declarations

Ethics approval and consent to participate

Confirmation was obtained from NRES that this was an evaluation of a service delivery intervention, and therefore neither ethical permission nor informed consent were required. Approval was obtained from each Research and Development Department.

Consent for publication

Not applicable

Availability of data and material

The datasets analysed during this study are available from the study statistician on reasonable request.

Competing interests

None

Funding

This work was funded by the National Institute for Health Research (NIHR) through the Collaborations for Leadership in Applied Health Research and Care for Birmingham and Black Country (CLAHRC-BBC) programme (CLAHRC 2013 10018). The views expressed are those of the authors and not necessarily those of the NHS, the NIHR or the Department of Health.

Authors' contributions

SK conceived the study with PC (Consultant Midwife, BWNFT) and AH (Associate Head of Midwifery, Heart of England Foundation Trust (HEFT)). SK and KH designed the study. SK was responsible for the day to day management and implementation supported by SD and LH (Research Midwife, UoB). KH is the study statistician, supported by DJ. SK drafted the paper and all authors have seen and approved the manuscript.

Acknowledgements

Thanks to Shahnaz Khan for data extraction; Sue Smithson, Wendy Burt, and Becky Wilson-Crelin, Viv Carbury who supported training delivery and Nicola Gale who co-facilitated the workshop for the midwives with Lucy Hope. Thanks to Kelly Hard (R&D Manager, BWNFT) and Liz Adey (R&D Manager, HEFT). Thank you to Rebecca Taylor and Ed Juszcak for helpful comments on earlier drafts of the paper.

REFERENCES

- ¹ NHS Maternity Statistics (2011) <http://www.ic.nhs.uk/statistics-and-data-collections/hospital-care/maternity>
- ² Shetty, A, Burt R, Templeton A. Women's perceptions, expectations and satisfaction with induced labour—A questionnaire-based study. *Eur J Obstet Gynecol Reprod Biol.* 2005;123:56-61.
- ³ National Institute for Health and Clinical Excellence. Inducing Labour. London NICE 2008. <http://guidance.nice.org.uk/CG70>
- ⁴ <http://www.nice.org.uk/guidance/cg70/documents/cg70-induction-of-labour-surveillance-review-decision-may-20142>
- ⁵ <http://www.nice.org.uk/guidance/qs22/chapter/Quality-statement-12-Fetal-wellbeing-membrane-sweeping-for-prolonged-pregnancy>
- ⁶ Adams G, Gulliford MC, Ukoumunne OC, Eldridge S, Chinn S, Campbell MJ. Patterns of intra-cluster correlation from primary care research to inform study design and analysis. *J Clin Epidemiol.* 2004;57(8):785-94.
- ⁷ Hussey MA, Hughes JP. Design and analysis of stepped wedge cluster randomized trials *Contemp Clin Trials.* 2007;28(2):182-91.
- ⁸ Hemming K, Lilford R, Girling AJ. Stepped-wedge cluster randomised controlled trials: a generic framework including parallel and multiple-level designs. *Stat Med.* 2015;34(2):181-96. doi: 10.1002/sim.6325. Epub 2014 Oct 24.
- ⁹ Hemming K, Girling A. A menu driven facility for sample size for power and detectable difference calculations in stepped wedge randomised trials. *STATA J.* 2014;14(2):363-80.
- ¹⁰ [Zou G. A Modified Poisson Regression Approach to Prospective Studies with Binary Data. Am J Epidemiol 2004; 159\(7\):702-6.](#)
- ¹¹ Kirkham M, Stapleton H. (2001) Informed Choice in Maternity Care: An Evaluation of Evidence Based Leaflets. Available from: https://www.york.ac.uk/inst/crd/CRD_Reports/crdreport20.pdf [Accessed 14/11/2014].
- ¹² Salas E, Cannon-Bowers J.A. The science of training: A decade of progress. *Annu Rev Psychol.* 2001;52:471-99.
- ¹³ Hemming K, Haines TP, Chilton PJ, Girling AJ, Lilford RJ. The stepped wedge cluster randomised trial: rationale, design, analysis, and reporting. *BMJ (Clinical research ed).* 2015;350:h391.
- ¹⁴ [Chen YF, Hemming K, Stevens AJ, Lilford RJ. Secular trends and evaluation of complex interventions: the rising tide phenomenon. BMJ Qual Saf. 2016 May;25\(5\):303-10. doi: 10.1136/bmjqs-2015-004372. PubMed PMID: 26442789; PubMed Central PMCID: PMC4853562.](#)
- ¹⁵ Michie S, van Stralen M, West R. The Behaviour Change Wheel: a new method for characterising and designing behaviour change interventions. *Implement Sci.* 2011;6:42.
- ¹⁶ Michie S. Implementation science: understanding behaviour change and maintenance. *BMC Health Serv Res.* 2014;14(Suppl 2):O9. doi:10.1186/1472-6963-14-S2-O9

Legend

Figure 1: Consort Flow Diagram

Table 1: Participant baseline characteristics

Table 2: Primary and secondary outcomes, sub-groups by Trust and process outcomes

Table 3: Overall evaluation of training

Figure S1: Trial design

Figure S2: Consort checklist for cluster trials

Figure S3: Demonstration of the underlying secular trend of the numbers of women having a membrane sweep over the weeks of the study

Table 1: Maternal characteristics

Table S1 – Comparison of the influence of maternal characteristics on sweeping before and after training

Influence of	Before training n=1417		After training n=1356		Relative Risk (95% CI)	P-value
Women’s age (years)						
<20 years	39/77	(50.6%)	40/81	(49.4%)	1.13 (0.57, 2.24)	0.74
20-35 years	488/1124	(43.4%)	486/1064	(45.7%)	0.79 (0.61, 1.03)	0.08
>35 years	102/216	(47.2%)	108/211	(51.2%)	1.11 (0.74, 1.67)	0.62
Parity						
Nulliparous	363/849	(42.8%)	345/788	(43.8%)	0.92 (0.68, 1.24)	0.60
Multiparous	266/568	(46.8%)	289/568	(50.9%)	0.88 (0.64, 1.21)	0.45
Ethnicity						
Africa	27/89	(30%)	33/81	(41%)	1.05 (0.49, 2.28)	0.90
Asia – South	151/442	(34%)	126/427	(30%)	0.76 (0.51, 1.15)	0.19
Asia – Other	5/16	(31%)	9/19	(47%)	0.80 (0.18, 3.58)	0.77
Caribbean	20/60	(33%)	21/57	(37%)	0.95 (0.37, 2.41)	0.91
European – Britain	359/642	(56%)	384/619	(62%)	0.99 (0.79, 1.24)	0.91
European – Other	24/65	(37%)	24/55	(44%)	1.06 (0.44, 2.56)	0.90
Middle East	11/43	(26%)	11/47	(23%)	1.24 (0.41, 3.74)	0.70
Other	32/60	(53%)	23/42	(55%)	1.11 (0.48, 2.54)	0.81
BMI						
≤18	25/57	(43.9%)	24/54	(44.4%)	0.84 (0.35, 2.00)	0.69
19-34	554/1251	(44.3%)	561/1189	(47.2%)	0.87 (0.68, 1.11)	0.25
≥35	50/108	(46.3%)	48/111	(43.2%)	0.99 (0.53, 1.86)	0.98
Index of multiple deprivation from postcode						
Quintile 1	364/908	(40.1%)	338/845	(40.0%)	0.90 (0.67, 1.20)	0.46
Quintile 2	121/244	(49.6%)	143/252	(56.7%)	0.88 (0.60, 1.30)	0.52
Quintile 3	102/191	(53.4%)	93/158	(58.9%)	1.02 (0.67, 1.54)	0.93
Quintile 4	31/52	(59.6%)	44/67	(65.7%)	1.29 (0.62, 2.66)	0.50
Quintile 5	11/20	(55.0%)	16/32	(50.0%)	1.20 (0.34, 4.24)	0.77

Analysis of outcomes excludes those with missing outcome data

Table 1 – Participant baseline characteristics

Note we exclude those ineligible for sweeping and those delivering within the training transition period. Percentages are of total, and include any women with missing data on that variable.

Baseline characteristics	Before training n= 1420		After training n=1367	
Women’s age (years)				
Median, IQR	29 [25,32]		29 [25,33]	
<20 years	78	(5.5%)	81	(5.9%)
20-35 years	1126	(79.3%)	1074	(78.6%)
>35 years	216	(15.2%)	212	(15.5%)
Parity				
Nulliparous	850	(59.9%)	793	(58.0%)
Multiparous	569	(40.1%)	574	(42.0%)
Ethnicity				
Africa	89	(6%)	81	(6%)
Asia – South	442	(31%)	427	(31%)
Asia – Other	16	(1%)	19	(1%)
Caribbean	60	(4%)	57	(4%)
European – Britain	642	(45%)	619	(45%)
European – Other	65	(5%)	55	(4%)
Middle East	43	(3%)	47	(3%)
Other	60	(4%)	42	(3%)
Unknown	3	(0%)	20	(1%)
BMI				
≤18	57	(4.0%)	54	(4.0%)
19-34	1254	(88.3%)	1199	(87.7%)
≥35	108	(7.6%)	112	(8.2%)
Index of multiple deprivation from postcode				
Quintile 1	911	(64.2%)	855	(62.5%)
Quintile 2	244	(17.2%)	253	(18.5%)
Quintile 3	191	(13.5%)	158	(11.6%)
Quintile 4	52	(3.7%)	67	(4.9%)
Quintile 5	20	(1.4%)	32	(2.3%)
Trust				
BWNFT	926	(65%)	871	(64%)
BHH	494	(35%)	496	(36%)

Table 3 – Overall evaluation of training

Evaluation of training	Low		Medium		High	
Knowledge of the evidence and current NICE Guidance regarding membrane sweeping						
Before training	1	2	3	4	5	
Immediately after training	1	2	3	4	5	
6 months after training	1	2	3	4	5	
Confidence in talking to a woman about membrane sweeping						
Before training	1	2	3	4	5	
Immediately after training	1	2	3	4	5	
6 months after training	1	2	3	4	5	
Knowledge of what is involved to enable you to undertake a membrane sweep						
Before training	1	2	3	4	5	
Immediately after training	1	2	3	4	5	
6 months after training	1	2	3	4	5	
Completion of the documentation regarding membrane sweeping and IOL						
6 months after training	1	2	3	4	5	
Did the training change your practice						
Yes	36 (60%)					
No	24 (40%)					
Has your champion been able to answer any questions you may have had						
Yes	33 (55%)					
No	2 (3%)					
I haven't asked any questions	22 (37%)					

Table 2 – Primary and secondary outcomes, sub-group by trust, and process outcomes

	Before training Number of women included =1417	After training Number of women included=1356	Relative Risk (95% CI)	P-value
Primary outcomes				
Women offered and accepting membrane sweeping *	629 (44.4%)	634 (46.8%)	0.90 (0.71, 1.13)	0.37
Mean average (SD) number of membrane sweeps per woman	0.603 (0.795)	0.627 (0.787)	RATE RATIO: 0.83 (0.67, 1.01)	0.068
Secondary outcomes				
Onset of labour**				
Induced	323 (22.8%)	328 (24.2%)	1.04 (0.80, 1.34)	0.77
Mode of birth***				
Instrumental	235 (16.6%)	233 (17.2%)	1.06 (0.75, 1.48)	0.75
Emergency CS	187 (13.2%)	177 (13.1%)	0.89 (0.63, 1.26)	0.52
Sub-group by Trust				
BWNFT – Adherence to Trust guidance				
All women swept at 40 weeks (39+4 – 40+3)	245/921 (26.6%)	253/868 (29.1%)	0.91 (0.64, 1.29)	0.596
All eligible women swept for a second time at 41 weeks**** (40+4 – 41+3)	78/504 (15.5%)	62/509 (12.2%)	0.75 (0.41, 1.36)	0.339
BHH – Adherence to Trust (NICE) guidance				
Nulliparous women swept at 40 weeks (39+4 – 40+3)	47/174 (27.0%)	59/173 (34.1%)	1.81 (0.84, 3.92)	0.131
All eligible nulliparous women swept for second time at 41 weeks*** (40+4 – 41+3)	10/80 (12.5%)	17/88 (19.3%)	2.28 (0.59, 8.87)	0.232
Multiparous women swept at 41 weeks (40+4 – 41+3)	38/152 (25.0%)	46/160 (28.8%)	0.78 (0.32, 1.88)	0.574
Process outcomes				
Sweeps offered but declined	80 (5.6%)	97 (7.2%)		
No record of sweeping	708 (50.0%)	625 (46.1%)		
Reason if abandoned				
Os closed	30 (4.8%)	28 (4.4%)		
Unable to reach	38 (6.0%)	20 (3.2%)		
Unable to sweep	50 (7.9%)	42 (6.6%)		
Other	13 (2.1%)	9 (1.4%)		
Location of sweep				
Community	400 (63.6%)	431 (68.0%)		

	Before training Number of women included =1417	After training Number of women included=1356	Relative Risk (95% CI)	P-value
Hospital	227 (36.1%)	195 (30.8%)		

RRs are estimated using a generalised linear mixed model and are adjusted for clustering and underlying temporal trends.

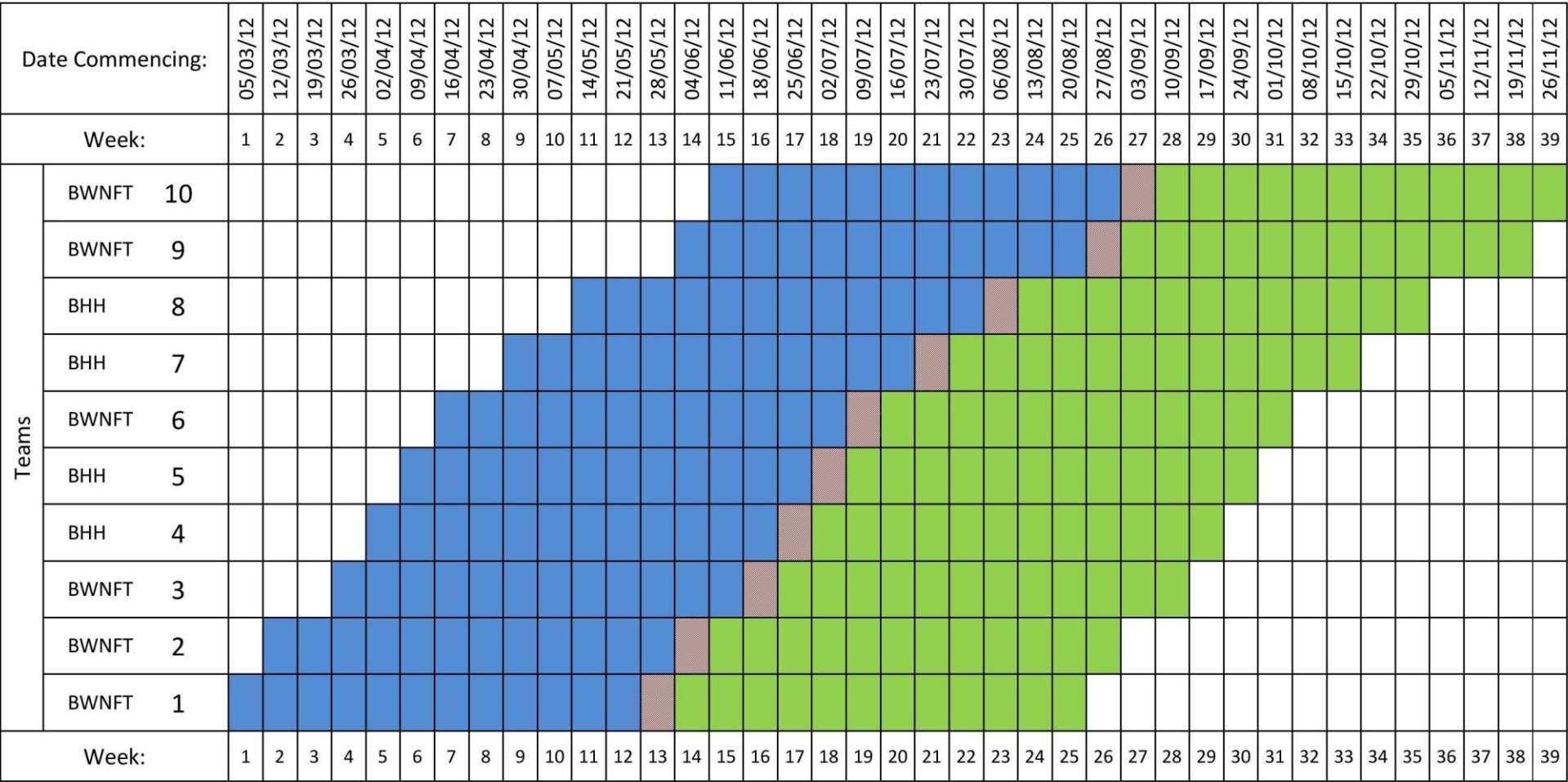
* The estimated ICC (95% CI) was 0.060 (0.000, 0.118) estimated using a one-way analysis of variance on the proportions scale.

** For onset of labour, the risk of being induced compared to spontaneous and not labouring combined was compared before and after training.

*** For mode of birth, the risk of instrumental birth compared to SVB and CS combined was compared before and after training. Separately, emergency CS was compared to SVB, instrumental and elective CS combined, before and after training.

**** Eligible women: pregnant at 41+3 weeks.

Supplementary Figure 1 - Stepped wedge cluster randomised trial design



Data were collected for 3 months before and after the point when the training was delivered in each team. Blue denotes the period of data collection prior to training, red shading denotes training, and green is the period of data collection following training.

Table 1: CONSORT 2010 checklist of information to include when reporting a cluster randomised trial

Section/Topic	Item No	Standard Checklist item	Extension for cluster designs	Page No *
Title and abstract				
	1a	Identification as a randomised trial in the title	Identification as a cluster randomised trial in the title	P1
	1b	Structured summary of trial design, methods, results, and conclusions (for specific guidance see CONSORT for abstracts) ^{1,2}	See table 2	P2
Introduction				
Background and objectives	2a	Scientific background and explanation of rationale	Rationale for using a cluster design	P5
	2b	Specific objectives or hypotheses	Whether objectives pertain to the the cluster level, the individual participant level or both	NA
Methods				
Trial design	3a	Description of trial design (such as parallel, factorial) including allocation ratio	Definition of cluster and description of how the design features apply to the clusters	P5
	3b	Important changes to methods after trial commencement (such as eligibility criteria), with reasons		NA
Participants	4a	Eligibility criteria for participants	Eligibility criteria for clusters	P5
	4b	Settings and locations where the data were collected		P7
Interventions	5	The interventions for each group with sufficient details to allow replication, including how and when they were actually administered	Whether interventions pertain to the cluster level, the individual participant level or both	P6
Outcomes	6a	Completely defined pre-specified primary and secondary outcome measures, including how and	Whether outcome measures pertain to the cluster level, the individual participant level or both	NA

when they were assessed				
	6b	Any changes to trial outcomes after the trial commenced, with reasons		NA
Sample size	7a	How sample size was determined	Method of calculation, number of clusters(s) (and whether equal or unequal cluster sizes are assumed), cluster size, a coefficient of intracluster correlation (ICC or k), and an indication of its uncertainty	P8-9
	7b	When applicable, explanation of any interim analyses and stopping guidelines		NA
Randomisation:				
Sequence generation	8a	Method used to generate the random allocation sequence		P6
	8b	Type of randomisation; details of any restriction (such as blocking and block size)	Details of stratification or matching if used	NA
Allocation concealment mechanism	9	Mechanism used to implement the random allocation sequence (such as sequentially numbered containers), describing any steps taken to conceal the sequence until interventions were assigned	Specification that allocation was based on clusters rather than individuals and whether allocation concealment (if any) was at the cluster level, the individual participant level or both	P8
Implementation	10	Who generated the random allocation sequence, who enrolled participants, and who assigned participants to interventions	Replace by 10a, 10b and 10c	
	10a		Who generated the random allocation sequence, who enrolled clusters, and who assigned clusters to interventions	P6
	10b		Mechanism by which individual participants were included in clusters for the purposes of the trial (such as complete	P7

enumeration, random sampling)				
	10c	From whom consent was sought (representatives of the cluster, or individual cluster members, or both), and whether consent was sought before or after randomisation		NA- service evaluation
Blinding	11a	If done, who was blinded after assignment to interventions (for example, participants, care providers, those assessing outcomes) and how		P8
	11b	If relevant, description of the similarity of interventions		
Statistical methods	12a	Statistical methods used to compare groups for primary and secondary outcomes	How clustering was taken into account	P9
	12b	Methods for additional analyses, such as subgroup analyses and adjusted analyses		P9
Results				
Participant flow (a diagram is strongly recommended)	13a	For each group, the numbers of participants who were randomly assigned, received intended treatment, and were analysed for the primary outcome	For each group, the numbers of clusters that were randomly assigned, received intended treatment, and were analysed for the primary outcome	P10
	13b	For each group, losses and exclusions after randomisation, together with reasons	For each group, losses and exclusions for both clusters and individual cluster members	Figure 1
Recruitment	14a	Dates defining the periods of recruitment and follow-up		P6
	14b	Why the trial ended or was stopped		NA
Baseline data	15	A table showing baseline demographic and clinical	Baseline characteristics for the individual and cluster levels as	Table 1

		characteristics for each group	applicable for each group	
Numbers analysed	16	For each group, number of participants (denominator) included in each analysis and whether the analysis was by original assigned groups	For each group, number of clusters included in each analysis	P10
Outcomes and estimation	17a	For each primary and secondary outcome, results for each group, and the estimated effect size and its precision (such as 95% confidence interval)	Results at the individual or cluster level as applicable and a coefficient of intracluster correlation (ICC or k) for each primary outcome	Table 2
	17b	For binary outcomes, presentation of both absolute and relative effect sizes is recommended		
Ancillary analyses	18	Results of any other analyses performed, including subgroup analyses and adjusted analyses, distinguishing pre-specified from exploratory		P11
Harms	19	All important harms or unintended effects in each group (for specific guidance see CONSORT for harms ³)		NA
Discussion				
Limitations	20	Trial limitations, addressing sources of potential bias, imprecision, and, if relevant, multiplicity of analyses		P12
Generalisability	21	Generalisability (external validity, applicability) of the trial findings	Generalisability to clusters and/or individual participants (as relevant)	P11-13
Interpretation	22	Interpretation consistent with results, balancing benefits and harms, and considering other relevant evidence		P11-13
Other information				
Registration	23	Registration number and		P3

		name of trial registry	
Protocol	24	Where the full trial protocol can be accessed, if available	NA
Funding	25	Sources of funding and other support (such as supply of drugs), role of funders	P16

** Note: page numbers optional depending on journal requirements*

Table 2: Extension of CONSORT for abstracts^{1,2} to reports of cluster randomised trials

Item	Standard Checklist item	Extension for cluster trials
Title	Identification of study as randomised	Identification of study as cluster randomised
Trial design	Description of the trial design (e.g. parallel, cluster, non-inferiority)	
Methods		
Participants	Eligibility criteria for participants and the settings where the data were collected	Eligibility criteria for clusters
Interventions	Interventions intended for each group	
Objective	Specific objective or hypothesis	Whether objective or hypothesis pertains to the cluster level, the individual participant level or both
Outcome	Clearly defined primary outcome for this report	Whether the primary outcome pertains to the cluster level, the individual participant level or both
Randomization	How participants were allocated to interventions	How clusters were allocated to interventions
Blinding (masking)	Whether or not participants, care givers, and those assessing the outcomes were blinded to group assignment	
Results		
Numbers randomized	Number of participants randomized to each group	Number of clusters randomized to each group
Recruitment	Trial status ¹	
Numbers analysed	Number of participants analysed in each group	Number of clusters analysed in each group
Outcome	For the primary outcome, a result for each group and the estimated effect size and its precision	Results at the cluster or individual participant level as applicable for each primary outcome
Harms	Important adverse events or side effects	
Conclusions	General interpretation of the results	
Trial registration	Registration number and name of trial register	
Funding	Source of funding	

¹ Relevant to Conference Abstracts

REFERENCES

- ¹ Hopewell S, Clarke M, Moher D, Wager E, Middleton P, Altman DG, et al. CONSORT for reporting randomised trials in journal and conference abstracts. *Lancet* 2008, 371:281-283
- ² Hopewell S, Clarke M, Moher D, Wager E, Middleton P, Altman DG at al (2008) CONSORT for reporting randomized controlled trials in journal and conference abstracts: explanation and elaboration. *PLoS Med* 5(1): e20
- ³ Ioannidis JP, Evans SJ, Gotzsche PC, O'Neill RT, Altman DG, Schulz K, Moher D. Better reporting of harms in randomized trials: an extension of the CONSORT statement. *Ann Intern Med* 2004; 141(10):781-788.

Figure S3 Demonstration of the underlying secular trend of the numbers of women having a membrane sweep over the weeks of the study

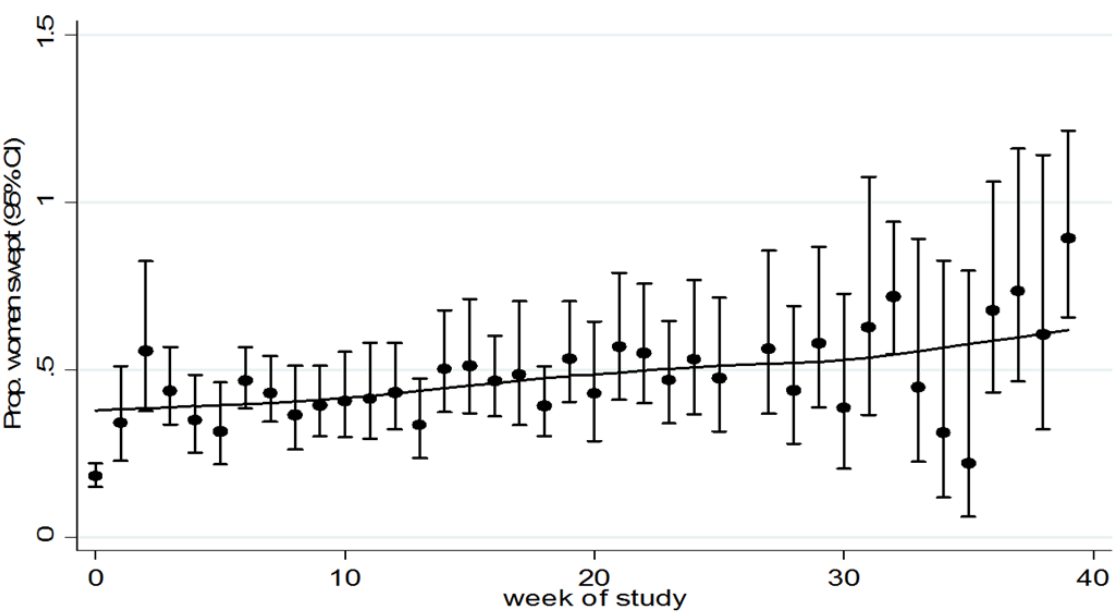


Figure 1 – CONSORT diagram

